International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Complex Event Processing of Health Data in Real Time Predicting Heart Failure Risk

Manaswini Sudhal, Deepak Gupta

***Abstract**: In this article, we develop a scalable system that can perform heart failure prediction techniques based on complex event processing (CEP). The emergence of different health conditions can be seen as complex events and therefore this concept can be easily extended to other uses. The system uses MLP (Multilayer Perceptron) for the prediction of heart failure. First, perform preprocessing and after that collect the health parameter. The system monitors the patients of heart failure and predicts heart attacks. When critical conditions are occurs the system warns the patients. Experimental results show that MLP is more accurate than C 4.5, based on Precision-Recall and F1.*

***Keywords** : Complex event processing, Heart Failures Prediction, C4.5, and Multilayer Perceptron.*

I. INTRODUCTION

Real-time health data collection is very common in todays and the ubiquitous variety of inexpensive health surveillance systems. These data can be exploited by a variety of signal processing and machine learning algorithms. The procedure for extraction and reasoning is similar in different applications. The various researchers working with real time signals and applying pre-processing step prior to making conclusions. The collected data can be used both in real-time and offline mode to draw several conclusions about the patient's condition [1]. However, in the application, the health sector is very limited due to the growing demand for networking and supporting infrastructure. For many users, real-world medical applications require the simultaneous analysis of high-resolution real-time sensor data and data from other sources. Local processing of all data on a single machine is impractical due to computational constraints, reliability, scalability, power issues, and recovery/ failures.

Analysis of health data, as a rule, is based on the comparison of health measures obtained on a predetermined threshold [2]. Symptoms can manifest when measuring above or below the threshold. Early detection of symptoms of heart failure enhances the prognosis of stroke and therefore prevents it. Therefore, the most important task is to determine the exact thresholds [3]. The accuracy of the analysis strictly depends on the accuracy of the used thresholds. Existing medical research, telemedicine is accepted, cardiac disease, diabetes, hypotension, hypertension, hyperthermia and hypothermia patients to the treatment [4]. The most promising program is a chronic disease of real-time monitoring, such as cardiopulmonary disease, asthma, and heart failure, health care facilities, from far away located patients in wireless surveillance systems.

Heart disease has become one of the leading causes of death in men worldwide. According to WHO, heart disease rates could increased worldwide by 23.3% up to 2030. This requires continuous and long-term monitoring to control the threat of the treatment of chronic diseases [1] [5]. Doctors present a new option for the user. We help wireless technology with a look at the possibilities for patient monitoring. The objective is to find solutions for Remote Patient Monitoring System Access.

Heart disease specialists define and update thresholds based on patient measurements and interviews conducted with the patient. The cardiologist's threshold is not the same for all patients and makes sure that the same may vary even for different patients [6]. As a result, health parameters are extracted and remotely monitoring method proposed will automatically calculate the run time health threshold to update an analytical method that will determine the heart failure patients.

II. REVIEW CRITERIA

Here present the literature review of existing techniques:

The paper demonstrates the health analysis technique for heart failure stroke prediction and uses a combination of CEP (complex event manipulation) techniques with a statistical approach for this benefit of this approach, is that it can automatically calculate the threshold value and update it upon execution. In this experiment, we used the MIMICII Waveform Database Matched subset of the physionet benchmark [1].

This paper presents a review of the literature on data mining and health management analysis using big data. Following PRISMA's guidelines (a systematic review and a priority reporting article for meta-analysis), they searched the database between 2005 and 2016. The essential elements of selected health sub-domains for research, technical data mining, type of analysis, data, and data sources are the areas and possible future they discover that the existing literature examines analysis primarily in clinical and administrative decision making [2].

Manuscript published on 30 September 2019
* Correspondence Author
Manaswini Sudhal*, Department of Computer Engineering, Siddhant College of Engineering Sadumbre, Pune, India. Email: smanup2001@gmail.com
Prof. Dr. Deepak Gupta*, Department of Computer Engineering, Siddhant College of Engineering Sadumbre, Pune, India. Email: deepak_gpt@yahoo.com

The amount of sensor data generated by each sensor is large; the storage of information becomes a problem. Manashty Alireza, Wright Janet and Yadav Umang [3], proposed to heal (Health Event Aggregation Lab) the proposed architecture of using previously processed similar data and associated identified symptoms are inputs sensors, devices and content providers for providing cloud services. The ultimate goal of the system is to bridge the gap between symptoms and diagnostic trend data to accurately and quickly predict health abnormalities.

The paper [4] proposed a system to develop a real-time cardiac monitoring system. The system is conceptualized to provide an interface for mutual communication between doctors and patients. The purpose of this article to help distant heart patients receive the latest in health care services that may not be possible otherwise due to the low doctor-to-patient ratio. Performance analysis shows that the proposed system is reliable and useful for its high speed.

The paper discusses WANDA (an activity with weight and blood pressure monitoring system) using sensors. This study is paired with the University of California School of nursing, UCLA Wireless Health Institute shows that CHF patients who readings monitored by Wanda are less likely to fall outside the healthy range to allow early detection of critical clinical symptoms that indicate a challenge associated with CHF. In addition, Wanda provides a useful feedback system to organize the reading of CHF patients [5].

In this article [6] personalized remote monitoring and evaluation of the user conditions. PhysioDroid System provides ubiquitous and continuous analysis of vital signs such as electrocardiogram, heart rate, respiration rate, and skin temperature and body movements to help patients become empowered and improve their clinical understanding. PhysioDroid consists of a portable monitoring device and Android application that provides the functionality for collection, storage, and processing of physiological sensors. Versatility of the application allows you to use it as secondary users and professionals, a reduction in the cost PhysioDroid makes it affordable for most people. Provides two examples of use for health assessment and sports coaching, to illustrate the capabilities of PhysioDroid’s.

Complex Event Processing (CEP) uses an event-driven approach and connects different sensor streams with spatiotemporal constraints for anomaly detection. This paper [6] proposes a CEP-based health monitoring system (CRHMS). The CRHMS proposed using biosensors (heart rate, respiratory rate, blood pressure, and ECG) to collect vital signs, and environmental sensors (accelerometer, global positioning system (GPS)) to identify the context of the elderly patient who is alone at home. These sensor settings are collected on an Android phone and sent as a stream to the proposed system to identify abnormalities of vital signs and generate an alarm.

According to various measures have emerged, and in recent years took part, namely, data stream processing systems (IPS). In this paper [7], they propose how semantic technology can contribute to the field of complex events and study their support in the field of health surveillance. Complex event processing system (CEP) combines precise semantic data with information being processed. The proposed approach combines Semantic Web way and CE model to a health monitoring platform.

Sensor data is an important mission of providing access to web theology, anytime anywhere. In particular, it is essential to monitoring the telemedicine of elderly patients who are alone at home. Continuous monitoring is necessary for the patient to wear wireless sensor data in the body for parameters. Sensor web Enablement data effectively and service-oriented architecture, the implementation of a clinical diagnosis, made for the physician to become available on the web raw sensor data to a platform. To know the complexity of threats, using the web to capture relationships between events, it is possible to overcome challenges [8].

System applications provide for activity records, events, and potentially important medical symptoms. The activity such as walking and running is detected from the movement of the body recorded by the accelerometer sensor which analyzes the T wave on the server where the features of the electrocardiogram are connected to the P wave of the electrocardiogram signal, the complex of QRS, and the base station which receives the data from the wireless sensor of the patient body. IEEE802. The 154 is used for wireless communication between the sensor and the base station. If an abnormality occurs on the server the alarm status sends to the doctor's personal digital assistant (PDA) [9].

The presence of a network of advances in the development of medical devices and the widespread use of data transmission allows you to equip more patients ‘interpretation of the collected data, resulting in the device is increasingly complex. Medical observation is traditionally interpreted in two competing ways, using established rule-based theories and statistically (which probably leads to a new theory). In this article, they will learn a hybrid approach that allows both the evaluation of fixed rule sets and their coexistence with machine learning [10].

III. METHODOLOGY

This section introduces Complex Event Processing of Health Data in Real Time Predicting Heart Failure Risk as a predictive analysis approach:

1) The system monitors the patients who are suffering from the disease.

2) Predicts heart failure strokes based on the related symptoms.

3) Warns to patients when the critical condition occurs.

Figure 1 shows the proposed system architecture of the system. Congestive heart failure (CHF) occurs when the heart cannot provide enough blood due to a healthy physiological state. CHF usually occurs when the cardiac tissue becomes chemical due to blockage of the coronary vessels. The proposed system includes modules are as follows: the processing Servers and Storage, data pre-processing, feature extraction, classification of heart defects using C 4.5 classifier and prediction of heart failure using multi-perceptron model (MLP). The system thresholds are calculated automatically using statistical approaches.

Thresholds are calculated based on the patient state and his/her historical measurements. The system main goal is to detect heart failure symptoms and predicting the critical conditions to warn the patients.

- The system firstly, stored data in a database such as CSV files. So, the history of patients is created.
- In the data preprocessing stage, data is preprocessed and missing values of replacing with threshold values or Zero.

- Once the preprocessing is done the system performs the Feature extraction process. The data feature values are calculated using principal component analysis methods. The goal of the PCA is to reduce data and rank columns with a high degree of impact. After selecting high impacted columns it removed characteristics from this column such as weight, blood pressure, body temperature.

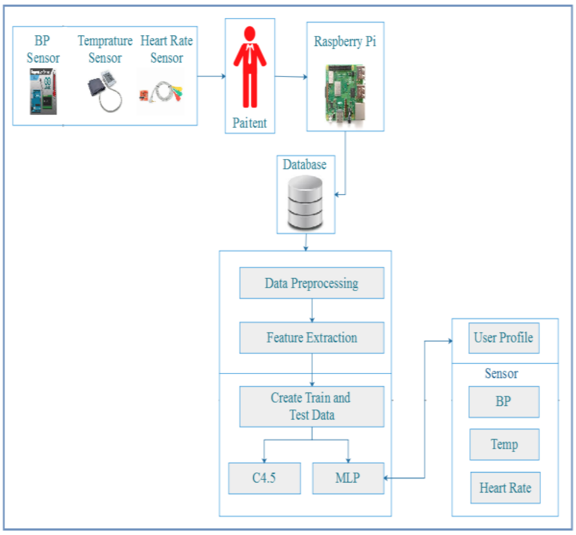

Fig. 1. System Architecture

- Then creating the training and testing files using extracted features and perform classification using C4.5 and MLP classifiers.
- Finally, this system recommends treatment to patients like Gym, stress level management.

A. Algorithm

- Algorithm 1: C4.5 Algorithm

1. Check for the below base cases:
   i. All the samples in the list belong to the same class. When this happens, it simply creates a leaf node for the decision tree saying to choose that class.
   ii. None of the features provide any information gain. In this case, C4.5 creates a decision node higher up the tree using the expected value of the class.
   iii. An instance of previously-unseen class encountered. Again, C4.5 creates a decision node higher up the tree using the expected value.
2. For each attribute, find the normalized information gain ratio from splitting on a.
3. Let a_best be the attribute with the highest normalized information gain.
4. Create a decision node that splits on a_best.
5. Recur on the sublists obtained by splitting on a_best, and add those nodes as children of the node.

- Algorithm 2: Multi-Level Perceptron Model (MLP)

Step 1: Initialize weights at random, choose a learning rate η

Until the network is trained:

For each training example i.e. input pattern and target output(s):

Step 2: Do forward pass through the net (with fixed weights) to produce output(s)

i.e., in forwarding Direction, layer by layer:

i. Inputs applied
ii. Multiplied by weights
iii. Summed
iv. 'Squashed' by the sigmoid activation function
v. Output passed to each neuron in the next layer

Repeat above until network output(s) produced

Step 3: Back-propagation of error

i. Compute error (delta or local gradient) for each output unit δ_k
ii. Layer-by-layer, compute error (delta or local gradient) for each hidden unit δ_j by back-propagating errors (as shown previously)

Step 4: Next, update all the weights Δwij By gradient descent, and go back to Step 2

The overall MLP learning algorithm, involving forward pass and backpropagation of error (until the network training completion), is known as the Generalized Delta Rule (GDR), or more commonly, the Back Propagation (BP) algorithm.

IV. RESULT AND DISCUSSION

A. Experimental Setup

The system is built by using Java framework and Netbeans IDE on windows platform. The system uses Heart Rate Sensor.

Heart Rate Sensor:

Parameter	Value
Operating Voltage:	+ 5V DC regulated
Operating Current:	100 mA
Output Data Level:	5V TTL level
Heart Beat detection:	Output High Pulse

B. Dataset

System use MIMIC II waveform database. The dataset contains around 33,000 patients of which approximately 25,000 are adults (having age ≥ 15 years old at the time of last admission) and around 8000 are neonates (age ≤ 1-month-old) at the time of the first admission. These patients experienced over 36,000 hospital admissions and over 40,000 ICU stays.

C. Results

In this section, we compare the existing system C4.5 algorithm and proposed system MLP algorithm. If it is proved that heart disease is present in the patient, then a given diagnostic test will also indicate the presence of heart disease, and the results of diagnostic tests will be considered true positive. Similarly, if the disease is not present in the patient, diagnostic tests suggest that heart disease is absent, and the result is a true negative (TN). A proven state of both true positive and true negative (also called the criterion of truth). But medical tests are not perfect.

If diagnostic tests show the presence of disease in patients who do not have such disease, the test results are false positive (FP). Similarly, if the results of the diagnostic examination suggest that the disease is definitely absent for patients with the disease, the test results are false-negative (FN). Both false positive and false negative indicate that the test result is opposite to the actual state.

Table- I: A Confusion Matrix

	Patient with Heart Disease	Patient without Heart Disease
Patient with Heart Disease	TP	FN
Patient without Heart Disease	FP	TN

We evaluated the result in terms of precision, recall and F1 Score. Precision, recall and F1 Score is calculated by using the following formulas:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

(2)

The accuracy in percent (%) of the system is calculated by using the formula:

$$\text{Accuracy} = \frac{\text{No. of correct classification of heart failure patient}}{\text{Number of all Patients}}$$

The system uses a standard dataset that contains 285 records. Out of 285 patients, the system has detected 258 patients suffering from heart failure. Table II shows the accuracy of the proposed system, in which the TP rate is 258, and it is precisely classified. 14 records are misclassified, FP rate is also 14 and FN rate is 14.

Table- II: Result for Accuracy of MLP Showing 98.99%

	Patient with Heart Disease	Patient without Heart Disease
Patient with Heart Disease	285	14
Patient without Heart Disease	14	10

Table- III: Comparison of C 4.5 and MLP Algorithm

Algorithm					
C4.5			*MLP*		
Precision	Recall	F1	Precisio n	Recall	F1
84.75%	100%	91.74 %	87.18%	98.99 %	95.00%

Figure 2 shows that the graph of the accuracy comparison between the existing system C 4.5 algorithm and the proposed system MLP algorithm. Our result shows that the proposed MLP archives 87.18 % precision, 98.99 % recall and 95.00 % F1 score. The result graph shows that the proposed system is more accurate than the existing system.

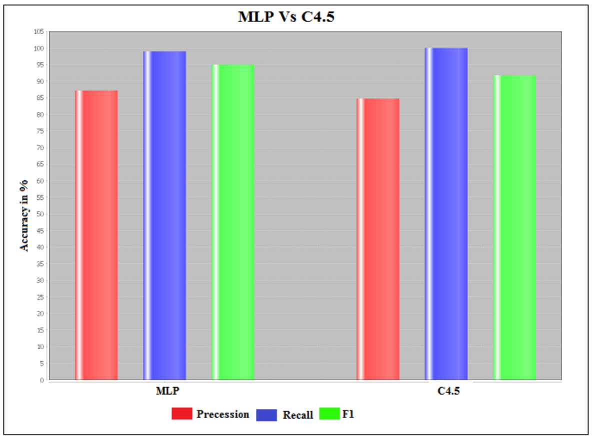

Fig. 2. Accuracy Comparison Graph between MLP and C 4.5 Algorithms

V. CONCLUSION

This paper presents a new approach to the prognostic analysis of cardiovascular diseases. It can be used to predict heart failure stroke. This system uses the CEP engine to continuously perform analysis of the collected health data. Based on the provisions on the analysis and some discussions with cardiologists, a general comparison represents the thresholds of the extracted health parameters. The system thresholds are calculated automatically using statistical approaches. Thresholds are calculated based on the patient state and his/her historical measurements. The experiment is carried out using MIMIC II waveform database used it matched subset from the Physionet benchmark. We evaluated the result in terms of precision, recall and F1 Score. Finally, Experimental results show that MLP is more accurate than C 4.5, based on Precision Recall and F1.

REFERENCES

1. G. O. Mdhaffar Afef, Charfi Khalil, Freisleben Bernd, And Abid Leila, “Cep4hfp: Complex Event Processing For Heart Failure Prediction”; Ieee Transactions On Nanobioscience, Vol. 16, No. 8, Ieee-[2017].
2. Md Saiful Islam, Md Mahmudul Hasan, Xiaoyi Wang , Hayley D. Germack, And Md Noor-E-Alam, “A Systematic Review On Healthcare Analytics: Application And Theoretical Perspective Of Data Mining”, Healthcare 2018, 6, 54; Doi:10.3390/Healthcare6020054, [2018].
3. Manashty Alireza, Light Janet, And Yadav Umang, "Healthcare Event Aggregation Lab (Heal), A Knowledge Sharing Platform For Anomaly Detection And Prediction"; Doi: 10.1109/Healthcom.2015.7454584, [2015].
4. P. Kakria, N. K. Tripathi, And P. Kitipawang, “A Real-Time Health Monitoring System For Remote Cardiac Patients Using Smart Phone And Wearable Sensors”, Int. J. Telemed. Appl., Vol. 2015, Art No. 8, [Jan. 2015].
5. Heinze Theodor, Wierschke Robert, Lowis Martin, And Schacht Alexander, “A Hybrid Artificial Intelligence System For Assistance In Remote Monitoring Of Heart Patients”; Doi: 10.1007/978-3-642-21222-2_50, Source: Dblp, [2014].
6. O. Banos, C. Villalonga, M. Damas, P. Gloesekoetter, H. Pomares, And I. Rojas, “Physiodroid: Combining Wearable Health Sensors And Mobile Devices For A Ubiquitous, Continuous, And Personal Monitoring,” Sci. World J., Vol. 2014, Art. No. 490824 [2014].

Retrieval Number: C6448098319/2019©BEIESP
DOI:10.35940/ijrte.C6448.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

7. Francesco Nocera, Tommaso Di Noia, Marina Mongiello, Eugenio Di Sciascio, "Knowledge-Enabled Complex Event Processing-Based Platform For Health Monitoring", Semapro 2016 : The Tenth International Conference On Advances In Semantic Processing, [2016].
8. Dhillon A.; Majumdar S.; El-Harak; And M. St-Hilaire, "Mcep: A Mobile Device Based Complex Event Processing System For Remote Healthcare"; Doi: 10.1109/Cybermatics_2018.2018.00064, Www.Spyglass-Consulting.Com, [2018].
9. Wang Chung, Punvar Amit, Bhardwaj Sachin, And Myllylae Esto, "A Fusion Health Monitoring Using Ecg And Accelerometer Sensors For Elderly Persons At Home"; Proceedings Of The 29th Annual International Conference Of The Ieee Embs Cite Internationale, Lyon, France, [2007].
10. Nocera Francesco, Mongiello Marina, And Sciascio Eugenio, "Knowledge-Enabled Complex Event Processing-Based Platform For Health Monitoring"; Semapro -The Tenth International Conference On Advances In Semantic Processing (Ticasp), [2016].

AUTHORS PROFILE

Manaswini Sudhal, P.G. Student, Department of Computer Engineering Savitribai Phule, Pune University Pune, Maharashtra

Dr. Deepak Gupta, B.E.,M.Tech (Hons),Ph.D(Computer Engineering), Associate Professor & HoD IT Siddhant College of Engineering Pune. Membership of LMISTE, MIAENG. 42 publications in International Journals and 9 publications in National Journal. 9 Conference Proceedings(Reviewed). 25 issued edited in different journals.